

Third-Order Phase-Locked Loop Perspectives

D. W. Brown

R.F. Systems Development Section

Analysis and practical application of third-order phase-lock design have been sporadic, compared with the second-order, in both the servomechanism and telecommunication fields. The attractiveness of minimal tracking errors resulting from near "perfect" third-order filtering (three true integrators) has been largely offset by undesirable acquisition properties and to some extent by a dearth of analysis of this configuration. A useful approach, both in viewpoint and in design, is to consider the prevalent "imperfect" second-order and third-order configurations for what they are—namely, loops with one integrator augmented by one or two lag time constants, so proportioned with respect to loop gain as to approximate the closed-loop response of true second- and third-order configurations, while manifesting a controlled (but not infinite) improvement in tracking performance over the first-order loop. This article seeks to apply this approach to the existing work in third-order analysis, and to emphasize the principal effects, both positive and negative, of the relative proportioning of loop gain and time constants, with a view toward practical exploitation of the best features of these loop configurations.

I. Introduction

Many works on phase-locked loops commence with first-order, progress to second-order, touch upon third-order, and then revert to an extensive treatment of second-order loops. While true that the latter is the most prevalent application, it seems clear that this sequence and depth of analysis are, in part, due to the fact that exact mathematical models exist for the first-order loop and quite adequate, but approximate, models exist for the second-order loop; and only tentative optimizations have appeared for the third-order loops. The word "tentative" is used here to imply that while there have existed theoretically optimum designs (Ref. 1) for the tracking of input

signals of specified characteristics, they may have lacked acceptable acquisition or stability properties over a range of input signal-to-noise ratios.

One of the more recent and detailed treatments (Ref. 2) of third-order loops appears to have made a significant step in overcoming these detrimental characteristics by evolving new optimization criteria. Whether by design or circumstance, this treatment by Tausworthe and Crow (Ref. 2) is closely structured to a "parallel" loop filter—that is, one which sums the output of a single-pole filter with that of a double-pole filter. This, coupled with the notation adopted, lends itself admirably to the switched second-to-

third order or "hybrid" application, as well as to a "straight" third-order design in which it is of interest to compare performance with that of the loop which results from the single-pole filter path only.

The purpose of this article is to apply the results and exactitude of the referenced work to a more conventional configuration by establishing an equivalence of the loop configurations and a somewhat re-oriented notation. This article will concern itself specifically with the optimum design and application of the third-order loop confronted with a variable signal level. Fortunately, this condition is virtually sufficient to establish the equivalence mentioned above. Therefore, this treatment will be equally applicable to both configurations. A further purpose is to summarize and extend the analysis to include optimum investment of the third-order characteristic and to establish a direct comparison with second-order tracking performance given the same circuit technology.

II. The General Transfer Function

Consider the phase-locked loop whose open-loop transfer function is

$$G(s) = \frac{G}{s} \frac{(1 + T_2 s)(1 + T_4 s)}{(1 + T_1 s)(1 + T_3 s)} \quad (1)$$

This loop has been considered (Ref. 3) with intuitive time constant scaling, but represents the form for which we seek a more rigorous and detailed analysis. Let us now proceed to develop a specific notation and optimization, borrowing heavily from Tausworthe (Refs. 2 and 4) and Hoffman (Ref. 5).

Tausworthe's parallel filter open-loop model is of the form

$$G'(s) = \frac{AK}{s} F(s) = \frac{AK}{s} \left[\frac{1 + \tau_2 s}{1 + \tau_1 s} + \frac{1}{(1 + \tau_1 s)(\delta + \tau_3 s)} \right] \quad (2)$$

with the following definitions:

$$\left. \begin{aligned} k &= \tau_2 / \tau_3 \\ \epsilon &= \tau_2 / \tau_1 \\ r &= AK\tau_2^2 / \tau_1 \\ AK &= \text{open-loop gain at zero frequency} \\ &\quad (\text{second-order path}) \\ \frac{AK}{\delta} &= \text{open-loop gain at zero frequency} \\ &\quad (\text{third-order path}) \end{aligned} \right\} \quad (3)$$

An optimum value for k has been established for which the loop will be unconditionally stable for all higher values of signal level, that is, for $r \geq r_0$ (design point). This value is $1/4$, subject to the restriction that the third-order lag time constant (τ_3/δ) be much larger than the lead compensation (τ_2) . This restriction is directly analogous to the usual second-order assumption of $\tau_2/\tau_1 = \epsilon \ll 1$.

III. Establishing Equivalence

Rewrite Eq. (2):

$$G'(s) = \frac{AK}{s} \left[\frac{1 + \tau_2 s}{1 + \tau_1 s} + \frac{1}{(1 + \tau_1 s)(\delta + \tau_3 s)} \right]$$

and, with a little manipulation,

$$G'(s) = \frac{AK}{s} \left(\frac{1 + \delta}{\delta} \right) \times \left[\frac{1 + \left(\frac{\delta}{1 + \delta} \right) (\tau_2 + \tau_3/\delta) s + \left(\frac{\delta}{1 + \delta} \right) \frac{\tau_2 \tau_3}{\delta} s^2}{(1 + \tau_1 s)(1 + \tau_3/\delta s)} \right] \quad (4)$$

The numerator is factorable with negative real roots, provided $\delta \ll 1$ and

$$\delta^2 \left(\tau_2 + \frac{\tau_3}{\delta} \right)^2 = 4\tau_2 \tau_3 \quad (5)$$

Examination of Eq. (4) reveals τ_1 and τ_3/δ to be the lag time constants, and τ_2 and τ_3 to be the approximate lead time constants and, therefore, δ represents the approximate lead-lag ratio, thus justifying the $\delta \ll 1$ approximation for moderate-to-high-performance loops. By similar reasoning,

$$\tau_2 + \frac{\tau_3}{\delta} \approx \frac{\tau_3}{\delta}$$

and Eq. (5) becomes

$$\tau_3^2 = 4\tau_2 \tau_3$$

$$\frac{\tau_2}{\tau_3} = k = \frac{1}{4}$$

Thus, we find the necessary condition for cascaded equivalence to be identical to the aforementioned criterion for stability versus signal level, namely, $k = 1/4$. A term-by-term comparison of Eqs. (1) and (4), that is, setting $G(s) = G'(s)$, now yields

$$\left. \begin{aligned}
 G &= AK \left(\frac{1 + \delta}{\delta} \right) \approx \frac{AK}{\delta} \\
 T_1 &= \tau_1 \\
 T_3 &= \frac{\tau_3}{\delta} \\
 T_2 T_4 &= \frac{\tau_2 \tau_3}{1 + \delta} \approx \tau_2 \tau_3 \\
 T_2 + T_4 &= \frac{\delta \tau_2 + \tau_3}{1 + \delta} \approx \tau_3
 \end{aligned} \right\} \quad (6)$$

and, setting $\tau_2/\tau_3 = 1/4$,

$$\begin{aligned}
 \frac{T_2 T_4}{(T_2 + T_4)^2} &= \frac{1}{4} \\
 T_2 = T_4 &= \frac{\tau_3}{2} = 2\tau_2
 \end{aligned}$$

The open-loop transfer function may now be written as

$$G(s) = \frac{G}{s} \left[\frac{(1 + T_2 s)^2}{(1 + T_1 s)(1 + T_3 s)} \right] \quad (7)$$

with relations (3) restated in terms of Eq. (7):

$$\left. \begin{aligned}
 \epsilon &= \frac{T_2}{2T_1} \\
 \frac{\delta}{4} &= \frac{T_2}{2T_3} \\
 r &= \frac{G}{2} \frac{T_2^3}{T_1 T_3}
 \end{aligned} \right\} \quad (8)$$

A word about approximations is perhaps in order at this point. The two measures of "imperfection" of the circuit "integrators" (ϵ and δk or $\delta/4$) will, for clarity, be taken as much less than unity, but never assumed to be zero. Indeed, the design "management" of these two small quantities (and related loop gain) constitutes a significant set of design tradeoffs. Therefore, only the first-order effects of ϵ and δ will usually be considered.¹

The definitions and equivalences of (3) and (6) are summarized in Appendix A. The identities and subsequent definitions based upon Eq. (7) are listed in Appendix B.

In order to establish the design point (minimum) value of the loop-gain-related variable, r , consider the closed-loop transfer function

$$H(s) = \frac{1 + 2T_2 s + T_2^2 s^2}{1 + \left(1 + \frac{1}{2GT_2}\right) 2T_2 s + \left(1 + \frac{T_1 + T_3}{GT_2}\right) T_2^2 s^2 + \frac{T_1 T_3}{G} s^3} \quad (9)$$

or

$$H(s) \approx \frac{1 + 2T_2 s + T_2^2 s^2}{1 + 2T_2 s + T_2^2 s^2 + \frac{T_1 T_3}{G} s^3}$$

Tausworthe's criterion of no underdamped roots in the interest of reliable acquisition can be applied by examining the characteristic equation, obtained by substitution:

$$s^3 + \frac{2r}{T_2} s^2 + \frac{4r}{T_2^2} s + \frac{2r}{T_2^3} = 0$$

A standard test for no imaginary roots reveals that

$$r \geq r_0 = \frac{27}{8} \quad (10)$$

As with the second-order loop, performance will be a function of r/r_0 , since

$$G = \frac{r}{r_0} G_0$$

¹For instance, the relations in (8) result from approximations in the establishment of equivalence and r could be written more exactly as

$$r = \frac{G}{2(1 + \delta)} \frac{T_2^3}{T_1 T_3}$$

On the other hand, one can take the view that once the equivalence and related optimizations have been established, relations in (8) become exact definitions for the cascade configuration. In fact, unpublished notes have *defined* a quantity

$$q = \frac{GT_2^3}{T_1 T_3}$$

yielding the equivalence relation

$$q = 2r(1 + \delta) \approx 2r$$

In the material that follows, G and r will usually be written as such for generality; however, $r_0 = 27/8$ will be implicitly assumed, especially in the next section dealing with approximations of comparative performance.

IV. Design Considerations

A. Definitions

We have now expended two of our five degrees of freedom²—that is, ratio of leads (function of k) and (design point) damping, $r = r_0$. Before investigating further the properties of the third-order loop, let us define:

$$\omega_1 = G \quad (11)$$

$$\omega_2 = \left(\frac{G}{T_1 + T_3} \right)^{1/2} \quad (12)$$

$$\omega_3 = \left(\frac{G}{T_1 T_3} \right)^{1/3} \quad (13)$$

By standard operations on Eq. (9), we see that ω_1 relates to (steady-state) phase error due to frequency offset and ω_2^2 to phase error due to a frequency ramp. Referring to Fig. 1, we see that ω_3 is the radian frequency at which the open-loop transfer function would pass through unity gain were it not for the lead compensation, T_2 . As such, ω_3 bears a close relationship to the bandwidth of the closed-loop response.

The distinctive advantage of the third-order over second-order loop lies in the ability to set ω_2 greater than ω_3 . In the second-order loop, ω_2 , or as usually written, $\omega_n = (G/\tau_1)^{1/2}$, represents both the closed-loop bandwidth and the reciprocal root of the ramp error coefficient.

B. Performance Factor

Let us define a performance factor, F , corresponding to the italicized statement above; that is,

$$F = \frac{\omega_2}{\omega_3}$$

$$F^6 = \left(\frac{G}{T_1 + T_3} \right)^3 \left(\frac{T_1 T_3}{G} \right)^2 = \frac{G T_1 \left(\frac{T_3}{T_1} \right)^2}{\left(1 + \frac{T_3}{T_1} \right)^3} \quad (14)$$

²Physically represented by G , T_1 , T_2 , T_3 , T_4 .

or, alternatively,

$$F^6 = \frac{G \sqrt{T_1 T_3} \left(\frac{T_3}{T_1} \right)^{3/2}}{\left(1 + \frac{T_3}{T_1} \right)^3}$$

$$F^2 = \frac{(G \sqrt{T_1 T_3})^{1/3} \left(\frac{T_3}{T_1} \right)^{1/2}}{\left(1 + \frac{T_3}{T_1} \right)} \quad (15)$$

We now find that given constant gain and bandwidth, and therefore constant $G \sqrt{T_1 T_3}$, F is maximized when

$$T_3 = T_1 \quad \text{or} \quad \frac{4\epsilon}{\delta} = 1$$

resulting in

$$F_{\text{opt}}^2 = \frac{(G T_1)^{1/3}}{2} = \frac{(2r)^{1/3}}{4\epsilon} \quad (16)$$

The conclusions here are that (1) for any given bandwidth and gain, the ramp error will be minimized when $T_1 = T_3$ (equal lags), and (2) the $G \sqrt{T_1 T_3}$ or $G T_1$ product is a measure of "how much third-order characteristic" is embodied in the design.

Now, let us generalize upon this result. As defined above, $F = \omega_2/\omega_3$ is a measure of third-order performance, but is based upon the $\epsilon, \delta \ll 1$ simplification. While this is a useful design assumption in many cases, a more general definition of F is

$$F_x = \frac{\text{Reciprocal root of ramp error coefficient}}{\text{Open-loop (undamped) unity gain frequency}} = \frac{\omega_2}{\omega_x}$$

The previously used expression for the numerator,

$$\omega_2 = \left[\frac{G}{T_1 + T_3} \right]^{1/2}$$

can be seen to be applicable over the full range of $0 \leq T_3/T_1 \leq \infty$. On the other hand, the denominator of F_x may be obtained by setting the absolute value of Eq. (7) to unity, yielding

$$\left(\frac{T_1 T_3}{G} \right)^2 \omega_x^6 + \left[\left(\frac{T_1 + T_3}{G} \right)^2 - \frac{2 T_1 T_3}{G^2} \right] \omega_x^4 + \frac{\omega_x^2}{G^2} = 1$$

The solution of this expression for ω_x is not conveniently expressible over the full range of T_3/T_1 . However, it will be stated without proof that whereas the unity gain frequency ω_x equals ω_3 for moderate ratios of T_3/T_1 , ω_x approaches $[G/T_1]^{1/2}$ or $[G/T_3]^{1/2}$ as T_3/T_1 approaches zero or infinity, respectively, and that

$$F_x^2 = \frac{\omega_2^2}{\omega_x^2} \approx 1 + \frac{\omega_2^2}{\omega_3^2} \quad (17)$$

Now consider a second-order loop of fixed G and T_1 , augmented by T_3 as T_3/T_1 increases from zero to infinity.

From Eqs. (14) and (17), we can now write

$$F_x^2 = 1 + \frac{(GT_1)^{1/2} \left(\frac{T_3}{T_1}\right)^{3/4}}{\left(1 + \frac{T_3}{T_1}\right)} \quad (18)$$

This expression is plotted with GT_1 as a parameter in Fig. 2. This presentation offers an insight to the effect of T_3 : As it increases (relative to T_1) from zero, it represents the typical "extra time constant" (Ref. 6) in a second-order loop until it becomes a stability "problem" in the transitional region, where it increases over $1/\omega_n$ to bring about a third-order characteristic. As T_3/T_1 eventually exceeds GT_1 , the loop reverts to one of second order, now with T_3 as the principal time constant and T_1 in the role of "extra."

The asymmetry of the curves results from the situation chosen for illustration, i.e., constant G and T_1 . If we now fix our attention near $T_3/T_1 = 1$ and ask: "How does the performance vary with T_3/T_1 , given a constant gain and closed-loop bandwidth," we get

$$F_x^2 = 1 + \frac{(G\sqrt{T_1 T_3})^{1/2} \left(\frac{T_3}{T_1}\right)^{1/4}}{\left(1 + \frac{T_3}{T_1}\right)} \quad (19)$$

as plotted in Fig. 3.

Consider now the time constant required by a second-order loop compared to a third-order design, again holding constant the gain and closed-loop bandwidth. By setting $\omega_n = \omega_3$,

$$\frac{G}{\tau_{12nd}} = \left(\frac{G}{T_1 T_3}\right)^{3/4}$$

$$\frac{\tau_{12nd}}{\sqrt{T_1 T_3}} = (G\sqrt{T_1 T_3})^{1/4}$$

or, if we set $T_3 = T_1$,

$$\frac{\tau_{12nd}}{T_1} = (GT_1)^{1/4} = 2F_{opt}^2$$

or, in terms of second-order measure,

$$\frac{\tau_{12nd}}{T_1} = (G\tau_1)^{3/4}$$

Since the GT_1 product is always a large number (see Fig. 1 and Appendix B), we find that the ω_2/ω_3 performance advantage is accompanied by a reduction in required time constant for the case under consideration ($\omega_n = \omega_3$ and $G = G$). The inverse statement of this effect (i.e., maintaining constant gain and time constant, third-order design offers narrower bandwidth) appears later in the section dealing with steady-state error.

It must be noted that because ω_3 has been chosen as representative of closed-loop bandwidth in the definition of F , Figs. 2 and 3 and the equations of this paragraph implicitly set $r = r_0 \ll F_{opt}$. For this reason, numerical application of this paragraph is not generally warranted for $G > G_0$.

C. Design Conclusions

The titles "Normalized ramp performance" for Figs. 2 and 3 arise through squaring the original definition of F_x above and by noting that for the second-order loop, $\omega_x = \omega_2 = \omega_n$. In other words, F_x^2 or ω_2^2/ω_x^2 may be considered as either the ratio of ramp performance for the two loop orders or, alternatively, as a measure of third-order ramp performance normalized to bandwidth.

Indeed, as F_x approaches unity, by definition, second-order performance prevails; similarly, as F_x approaches infinity, an "ideal" third-order loop is manifested, with infinitesimal steady-state errors. But extremely large values of loop gain and/or time constant are neither practically achievable nor necessarily desirable from an overall performance standpoint.

In summary, then, with or without the exact optimization of $T_1 = T_3$ (it is difficult to envision a circumstance where this should not be applied), the fifth degree of freedom³ consists of "modulating" the GT_1 product with respect to bandwidth selection in achieving a compromise

³As noted earlier, the ratio of leads and damping have been considered established; we may now consider T_1 versus T_3 and closed-loop bandwidth the third and fourth degrees, respectively.

between steady-state tracking errors and problems of design and performance associated with high GT_1 . Some of these performance characteristics will be examined in the next section.

V. Performance Characteristics⁴

A. Closed-Loop Frequency Response

Referring back to the approximate form of Eq. (9), and to the design point value of r given in Eq. (10), we may write

$$H(j\omega) = \frac{1 - \omega^2 T_2^2 + j2\omega T_2}{1 - \omega^2 T_2^2 + j\left(2\omega T_2 - \frac{4}{27}\omega^3 T_2^3\right)} \quad (20)$$

and the loop error

$$1 - H(j\omega) = \frac{-j\frac{4}{27}\omega^3 T_2^3}{1 - \omega^2 T_2^2 + j\left(2\omega T_2 - \frac{4}{27}\omega^3 T_2^3\right)} \quad (21)$$

The absolute values of these functions have been plotted in Fig. 4 with respect to the normalized ω/ω_3 for a third-order loop in comparison to the familiar second-order curves given as a function of ω/ω_n .

B. Loop Noise Bandwidth⁵

It has been shown that

$$2\beta_L = W_L = \frac{r}{2\tau_2} \left[\frac{r - k + 1}{r - k} \right]$$

which may be rewritten in the cascade notation as

$$2\beta_L = W_L = \frac{r}{T_2} \left(\frac{4r + 3}{4r - 1} \right) \quad (22)$$

$$2\beta_{L_0} = W_{L_0} = \frac{891}{200T_2}, \quad \text{given } r = r_0 = \frac{27}{8}$$

or, in terms of $\omega_3 = (G/T_1 T_3)^{1/3}$:

$$2\beta_L = W_L = \left(\frac{r^2}{2} \right)^{1/3} \left(\frac{G}{T_1 T_3} \right)^{1/3} \left(\frac{4r + 3}{4r - 1} \right) \quad (23)$$

$$2\beta_{L_0} = W_{L_0} = \frac{297}{100} \frac{\omega_{3_0}}{(2)^{1/3}} \approx 2.36 \omega_{3_0}$$

⁴Unless otherwise specified, ϵ and $\delta/4$ are assumed to be much less than unity, but greater than zero. Also, the equivalences of Eq. (6) apply throughout.

⁵ β_L and W_L are in hertz; all other frequencies are in radians/second.

For the second-order loop, we have

$$2\beta_L = W_L = \frac{r + 1}{2\tau_2} \quad (24)$$

$$2\beta_L = W_L = \frac{r + 1}{2\sqrt{r}} \left(\frac{G}{\tau_1} \right)^{1/2} \quad (25)$$

$$2\beta_{L_0} = W_{L_0} = \frac{3}{2\tau_2} \approx 1.06 \omega_{n_0}, \quad \text{given } r = r_0 = 2$$

C. Steady-State Error

The steady-state error in response to an input of $\Omega(t)$ instantaneous frequency offset and Λ_0 rate of change of frequency is given by

$$\Phi_{ss} = \frac{\Omega(t)}{G} + \Lambda_0 \frac{T_1 + T_3}{G}$$

or, in the case of $T_3 = T_1$,

$$\Phi_{ss} = \frac{\Omega(t)}{G} + \Lambda_0 \frac{2T_1}{G} = \frac{\Omega(t)}{G} + \frac{\Lambda_0}{\omega_n^2} \quad (26)$$

For the second-order loop,

$$\Phi_{ss} = \frac{\Omega(t)}{G} + \Lambda_0 \frac{\tau_1}{G} = \frac{\Omega(t)}{G} + \frac{\Lambda_0}{\omega_n^2} \quad (27)$$

Comparing these two expressions, we can formulate another second-to-third-order loop comparison as follows. Given a second-order loop design of G , τ_1 and appropriate damping, a third-order loop may be formed by halving the τ_1 filter and placing the halves in series as T_1 and T_3 ,⁶ maintaining the same G and readjusting the damping; the resulting third-order loop will have exactly the same steady-state phase error due to offset and ramp as the given loop and will have a noise bandwidth approximately $1/F_{\text{opt}}$ times narrower.

D. Acquisition Range

For the ideal (no hardware biases or nonlinearities) noiseless case, it has been shown that the maximum frequency pull-in range is given by

$$\Omega_0 \leq \frac{r}{\tau_2} \left[\frac{2\tau_1}{\tau_2} \left(\frac{1 + \delta}{\delta} \right) \right]^{1/2}$$

⁶Conservation of time constant $T_1 + T_3 = \tau_1$, while probably not a usual design concept, is of economic as well as mathematical consequence.

which, for third-order, $\delta \ll 1$ and $T_1 = T_3$, reduces to

$$\Omega_0 \leq \frac{\sqrt{2}r}{T_2\epsilon} = \frac{2\sqrt{r}}{T_2} \sqrt{GT_2} \quad (28)$$

and, for second-order, $\delta \gg 1$, yields

$$\Omega_0 \leq \frac{\sqrt{2}r}{\tau_2\sqrt{\epsilon}} = \frac{\sqrt{2}r}{\tau_2} \sqrt{G\tau_2} \quad (29)$$

In terms of gain (G) and bandwidth ($1/\tau_2$ or $1/T_2$), we find that pull-in ranges for the second- and third-order loops are comparable and yet there is a difference of $\epsilon^{1/2}$ in terms of bandwidth only ($1/\tau_2$ or $1/T_2$). This arises from the fact that for the $T_1 = T_3$ third-order loop, G is proportional to ϵ^{-2} , whereas for the second-order, G varies as ϵ^{-1} . This effect can also be observed for steady-state error. That is, for a given bandwidth, second-order step error varies as ϵ^{-1} and ramp error as ϵ^0 , while third-order step error varies as ϵ^{-2} and ramp error as $\epsilon^{-1/2}$, even though Eqs. (26) and (27) are essentially identical as expressed in terms of G and τ_1 or T_1 . Obviously, extreme care must be exercised in the drawing of comparative conclusions. These relationships result directly through application of the identities of Appendix B. Appendix C attempts to catalog these relationships by relating the orders-of-magnitude, in terms of ϵ and of G , of second- and third-order performance.

Returning to acquisition, unpublished work of Tausworthe shows that in the presence of hardware drift referred to phase error, θ_d , expressed in radians, pull-in from one side is not simply reduced in range in proportion to $\theta_d G$, but is conditional upon initial conditions of the acquisition process, unless

$$|\theta_d| < \sqrt{2}\epsilon$$

for the $T_1 = T_3$ third-order loop. So, while for negligible θ_d , pull-in range will be enhanced as $\epsilon, \delta \rightarrow 0$, the existence of a finite θ_d will bound the useful ϵ, δ unless circumstances of acquisition (initial conditions) are well controlled.

E. Design Limitations

We have found that, in general, performance characteristics of tracking loops are all enhanced as $\epsilon, \delta \rightarrow 0$, either in the secondary sense of validation of approximations or in the primary sense of steady-state tracking errors. However, in the paragraph above, we encountered a limitation. Other considerations of practical implementations which are negative attributes of high-gain designs include:

- (1) Cost and resistance shunting (leakage) of large capacitors and of extremely high-gain operational amplifier configurations.
- (2) Unpredictability of performance under conditions of leakage versus time, environment, and hardware sample. This can be partially overcome through the use of compensating circuit configurations, such as with capacitance in an operational feedback configuration, gain and time constant are equally affected by shunting, thus yielding a constant G/T ratio and consequent stabilization of some aspects of performance.
- (3) Design difficulties related to saturation of operational amplifiers and VCOs.
- (4) Acquisition uncertainties due to input noise and hardware drift.
- (5) Operational limitations resulting from combinations of the above.

Since these limitations are so dependent upon specific circuit technologies and configurations, as well as particular performance applications, it would be virtually impossible to offer any comprehensive analytical treatment of these effects. Rather, the approach here has been an attempt to clarify the performance through the perspectives and relationships established earlier and summarized in Appendix C, thus allowing the designer to choose an appropriate loop order and gain to just satisfy his performance objectives while realizing a tolerable set of "negative" effects.

F. Acquisition Strategies

While considering closed-loop (tracking) performance, it may have been noted that the two extra degrees of freedom (T_3 and T_4) available in third-order loop design were spent early, in the interest of stability and minimization of ramp error ($T_1 = T_2$ and $T_3 = T_4$). To put it more directly, third-order design and second-order design both reduce to a choice of loop gain and bandwidth. This rationale assumes that the established values of design point damping are beyond question for the variable signal level tracking loop. Actually, for the third-order design, the criterion of critical damping at design point ($r_0 = 27/8$), in the interest of reliable acquisition, may be a bit conservative. As with the second-order practice of $r_0 = 2$, one could design for a slight underdamping at design point and still realize unconditional stability at all higher signal levels. While rigorously related mathematically (see Appendices A and B and Refs. 2 and 4), the significance of a given value of r_0 in each loop order is somewhat different. It has been pointed out that more analysis in this area may be enlightening.

Assuming now that all the foregoing has established a tracking design, there are two commonly used strategies of acquisition taken together or separately that alter the loop (additional external aids such as frequency sweeping will not be considered here). Firstly, one can seek to control the initial conditions—commonly implemented and referred to as either “open” or “shorted” loop. Extra care must be exercised here for third-order design due to the existence of twice as many repositories for initial charge as in the more familiar second-order design. And secondly, one may choose to modify one or more of the five loop parameters to enhance pull-in range or time and/or to reduce (until lock is achieved) the effect of a hardware imperfection. Several strategies that have been used in second-order applications are:

- (1) Reduce τ_1 and τ_2 , constant gain and ϵ .
- (2) Reduce τ_1 , constant gain and τ_2 .
- (3) Reduce τ_1 and τ_2 , constant gain and r .

The first reduces acquisition time but doesn't affect acquisition range; the second does both, and the third is a completely new bandwidth design, etc.

Now, to consider additional possibilities for the third order, a switched second-to-third-order loop strategy is possible; also, gain switching with or without time constant switching has been utilized to some extent. To illustrate the complexity of the problem, one can readily imagine a circumstance wherein one would face a dilemma in loop gain: Considerations of pull-in time or range might seem to demand an increased gain during acquisition and

yet this might invoke problems related to initial conditions and equipment drift (θ_d). The answer in a given situation might lie in switching one or more time constants in addition to or instead of loop gain. Here, again, the relationships of Appendix C may provide the insight to achieve a workable strategy. The significant point is that, with several degrees of freedom available (criteria established for optimum tracking performance are not necessarily applicable during acquisition), the difference, for instance, in a linear and a square root dependence may offer the answer to a given situation.

VI. Concluding Remarks

Partial conclusions have been drawn as the discussion and analysis proceeded above. Suffice it to say here that it is hoped that this treatment has provided a “feel” for the characteristics and relationships of the third-order loop, not only in an absolute sense, but in relation to second-order characteristics. And, perhaps most importantly, some of the stigma of operational unreliability of this device has hopefully been removed through demonstration that a number of undesirable acquisition and design characteristics tend to result not from the loop order but from the application of excessive loop gain in the guise of more “perfect” low-pass “integrators.”

In summary, third-order design can offer (other things equal) reduced time constant requirement, narrower closed-loop bandwidth, reduced tracking errors or combinations of these with slight increase in complexity over typical high-performance second-order designs.

References

1. Viterbi, A. J., *Principles of Coherent Communication*, p. 66. McGraw-Hill Book Co., Inc., New York, 1966.
2. Tausworthe, R. C., and Crow, R. B., *Practical Design of Third-Order Phase-Locked Loops*, Report 900-450, Apr. 27, 1971 (JPL internal document).
3. Gardner, F. M., *Phaselock Techniques*, p. 15. John Wiley & Sons, Inc., New York, 1966.
4. Tausworthe, R. C., *Theory and Practical Design of Phase-Locked Receivers*, Technical Report 32-819. Jet Propulsion Laboratory, Pasadena, Calif., Feb. 15, 1966.
5. Hoffman, L. A., *Receiver Design and the Phase-Lock Loop*. The Aerospace Corporation, El Segundo, Calif., May 1963.
6. Rechtin, E., *Design of Phase-Lock Oscillator Circuits*, Section Report No. 8-566, Feb. 7, 1957 (JPL internal document).

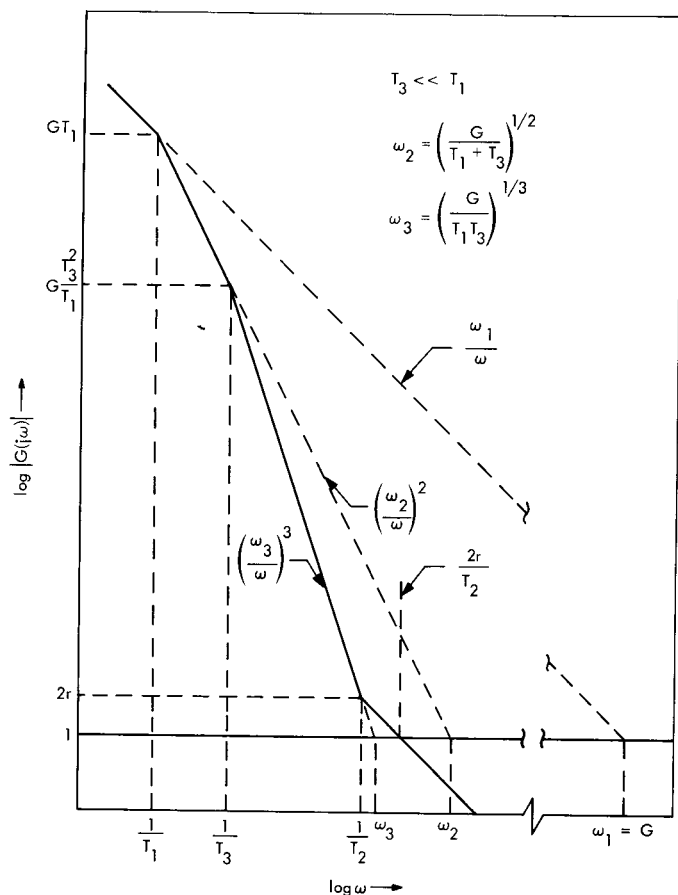


Fig. 1. Idealized open-loop frequency response

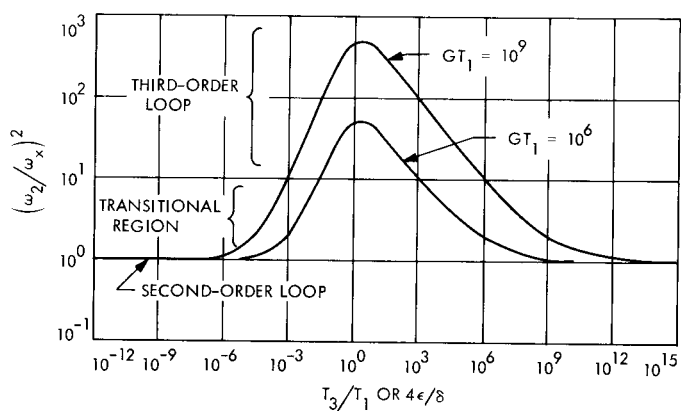


Fig. 2. Normalized ramp performance (fixed G and T_1)

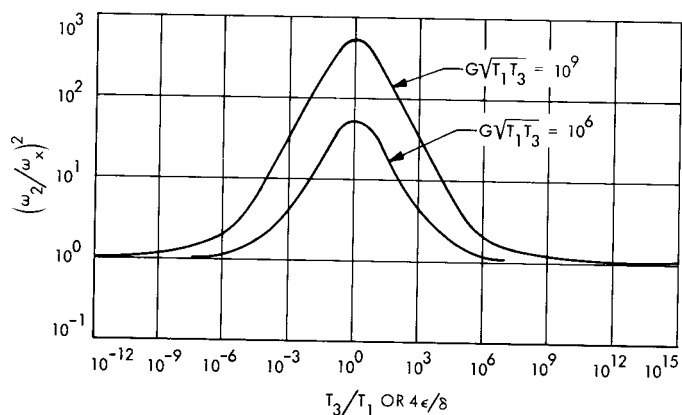


Fig. 3. Normalized ramp performance (fixed G and bandwidth)

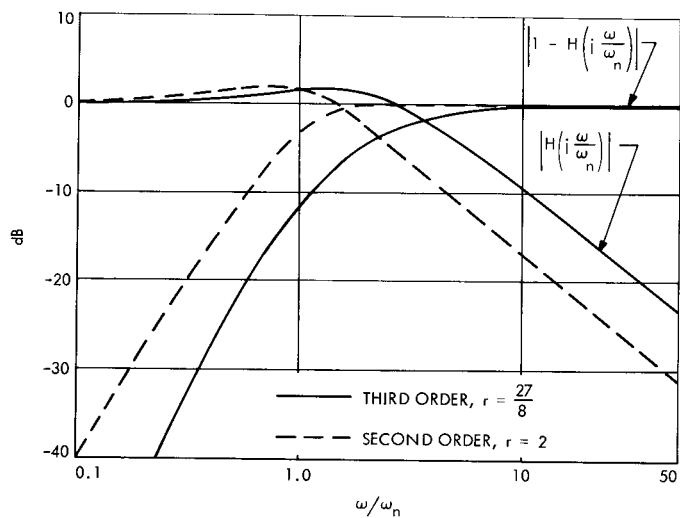


Fig. 4. Closed-loop frequency response

Appendix A

Summary of Definitions and Equivalences

Definitions (from Refs. 2 and 4)

$$k = \tau_2/\tau_3$$

$$\epsilon = \tau_2/\tau_1$$

$$r = AK\tau_2^2/\tau_1$$

AK = Open-loop gain of second-order portion of
“parallel” configuration at zero frequency

$\frac{AK}{\delta}$ = Open-loop gain of third-order portion of
“parallel” configuration at zero frequency

Equivalences (given $k = 1/4$ and $\delta \ll 1$)

$$G = AK \left(\frac{1 + \delta}{\delta} \right) \approx \frac{AK}{\delta}$$

$$T_1 = \tau_1$$

$$T_2 \approx \frac{\tau_3}{2} = 2\tau_2$$

$$T_3 = \frac{\tau_3}{\delta}$$

$$T_4 = T_2$$

Appendix B

Summary of Derived Identities and Definitions

Basic Identities

$$\epsilon = \frac{T_2}{2T_1}$$

$$\delta = \frac{2T_2}{T_3}$$

$$r = \frac{G}{2} \frac{T_2^3}{T_1 T_3}$$

Derived Identities

$$\text{General} \quad T_3 = T_1$$

$$\frac{T_3}{T_1} = \frac{4\epsilon}{\delta} \longrightarrow 1$$

$$GT_1 = \frac{r}{\delta\epsilon^2} \longrightarrow \frac{r}{4\epsilon^3}$$

$$GT_2 = \frac{2r}{\delta\epsilon} \longrightarrow \frac{r}{2\epsilon^2}$$

$$\frac{2r}{T_2} = \epsilon\delta G \longrightarrow 4\epsilon^2 G$$

$$G\sqrt{T_1 T_3} = \frac{2r}{(\delta\epsilon)^{3/2}} \longrightarrow \frac{r}{4\epsilon^3}$$

Supplemental Definitions

$$\omega_1 = G = \text{Reciprocal step (frequency) error coefficient}$$

$$\omega_2 = \left[\frac{G}{T_1 + T_3} \right]^{1/2} = \text{Reciprocal square root of ramp (frequency) error coefficient}$$

$$\omega_3 = \left[\frac{G}{T_1 T_3} \right]^{1/3} = \text{Natural (undamped) frequency}$$

$$\omega_n = \left[\frac{G}{\tau_1} \right]^{1/2} = \text{Natural (undamped) frequency and reciprocal square root of ramp (frequency) error coefficient for second-order loop}$$

$$\omega_r = \text{Open loop (undamped) unity gain frequency; equal to } \omega_3 \text{ for third-order and to } \omega_2 \text{ or } \omega_n \text{ for second-order loop}$$

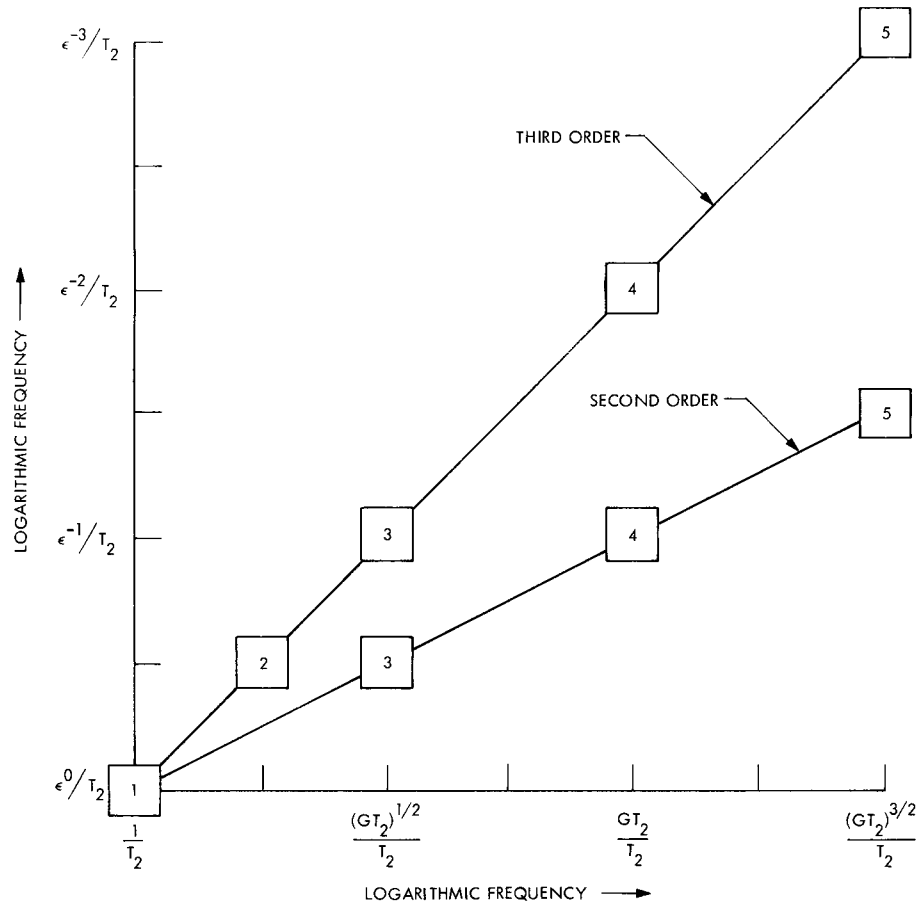
$$F = \frac{\omega_2}{\omega_3} = \text{A performance factor relating ramp error to closed-loop bandwidth}$$

$$F_r = \frac{\omega_2}{\omega_r} = \text{A generalization of } F \text{ for any } \delta \text{ from zero to infinity, holding } \epsilon \ll 1$$

Appendix C

Characteristics Versus G Versus Gain, Given $T_1 = T_3$, $1 < r < 10$

This figure is perhaps best interpreted as an order-of-magnitude conversion chart for ϵ and G for each loop order. The $T_2 \approx \tau_2$ normalization and notes allow superposition of the principal characteristics.



- 1 BANDWIDTH, BOTH ORDERS; RECIPROCAL ROOT RAMP COEFFICIENT, SECOND ORDER
- 2 RECIPROCAL ROOT RAMP COEFFICIENT, THIRD ORDER
- 3 MAXIMUM PULL-IN RANGE, BOTH ORDERS
- 4 RECIPROCAL STEP ERROR COEFFICIENT, BOTH ORDERS
- 5 $(GT_2)^{3/2} \approx GT_1$ (THIRD ORDER) $\approx (GT_1)^{3/4}$ (SECOND ORDER)